JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY
© 2015 BY THE AMERICAN COLLEGE OF CARDIOLOGY FOUNDATION
PUBLISHED BY ELSEVIER INC.

VOL. 66, NO. 15, 2015
ISSN 0735-1097/$36.00
http://dx.doi.org/10.1016/j.jacc.2015.08.035

ORIGINAL INVESTIGATIONS

10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors

Derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) With Validation in the HNR (Heinz Nixdorf Recall) Study and the DHS (Dallas Heart Study)

Robyn L. McClelland, PhD,* Neal W. Jorgensen, MS,* Matthew Budoff, MD,† Michael J. Blaha, MD, MPH,‡ Wendy S. Post, MD, MS,‡ Richard A. Kronmal, PhD,* Diane E. Bild, MD, MPH,§ Steven Shea, MD, MS,‖ Kiang Liu, PhD,¶ Karol E. Watson, MD, PhD,# Aaron R. Folsom, MD,** Amit Khera, MD,†† Colby Ayers, MS,‡‡ Amir-Abbas Mahabadi, MD,§§ Nils Lehmann, PhD,‖‖ Karl-Heinz Jöckel, PhD,‖‖ Susanne Moebus, PhD,‖‖ J. Jeffrey Carr, MD, MS,¶¶ Raimund Erbel, MD, PhD,§§ Gregory L. Burke, MD, MS##

ABSTRACT

BACKGROUND Several studies have demonstrated the tremendous potential of using coronary artery calcium (CAC) in addition to traditional risk factors for coronary heart disease (CHD) risk prediction. However, to date, no risk score incorporating CAC has been developed.

OBJECTIVES The goal of this study was to derive and validate a novel risk score to estimate 10-year CHD risk using CAC and traditional risk factors.

METHODS Algorithm development was conducted in the MESA (Multi-Ethnic Study of Atherosclerosis), a prospective community-based cohort study of 6,814 participants age 45 to 84 years, who were free of clinical heart disease at baseline and followed for 10 years. MESA is sex balanced and included 39% non-Hispanic whites, 12% Chinese Americans, 28% African Americans, and 22% Hispanic Americans. External validation was conducted in the HNR (Heinz Nixdorf Recall Study) and the DHS (Dallas Heart Study).

RESULTS Inclusion of CAC in the MESA risk score offered significant improvements in risk prediction (C-statistic 0.80 vs. 0.75; $p < 0.0001$). External validation in both the HNR and DHS studies provided evidence of very good discrimination and calibration. Harrell's C-statistic was 0.779 in HNR and 0.816 in DHS. Additionally, the difference in estimated 10-year risk between events and nonevents was approximately 8% to 9%, indicating excellent discrimination. Mean calibration, or calibration-in-the-large, was excellent for both studies, with average predicted 10-year risk within one-half of a percent of the observed event rate.

CONCLUSIONS An accurate estimate of 10-year CHD risk can be obtained using traditional risk factors and CAC. The MESA risk score, which is available online on the MESA web site for easy use, can be used to aid clinicians when communicating risk to patients and when determining risk-based treatment strategies. (J Am Coll Cardiol 2015;66:1643–53)

From the *Department of Biostatistics, University of Washington, Seattle, Washington; †Division of Cardiology, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Los Angeles, California; ‡Johns Hopkins Ciccarone Center for the Prevention of Heart Disease, Baltimore, Maryland; §Patient-Centered Outcomes Research Institute, Washington, DC; ‖Departments of Medicine and Epidemiology, Columbia University, New York, New York; ¶Department of Preventive Medicine,

ABBREVIATIONS AND ACRONYMS

CAC = coronary artery calcium
CHD = coronary heart disease
ECG = electrocardiogram
MI = myocardial infarction

Coronary artery calcium (CAC) scores derived from routine cardiac-gated noncontrast computed tomography scans are a commonly used method for enhancing clinical cardiovascular risk prediction. Importantly, CAC scores are incremental but not redundant with traditional risk factors, and therefore, integration of both sets of information can enhance risk assessment. Indeed, the added value of CAC over and above traditional risk factors for prediction of cardiovascular events has been demonstrated in several studies (1–11). However, to date, no published risk scores are available to clinicians to incorporate CAC into routine 10-year risk prediction.

SEE PAGES 1654 AND 1669

The MESA (Multi-Ethnic Study of Atherosclerosis), due to its population-based, multiethnic composition and availability of 10 years of follow-up for incident CHD events, provides a unique opportunity to describe how CAC might be optimally combined with traditional risk factors in risk prediction. In this paper, we describe a novel MESA risk score that can be used to estimate 10-year CHD risk in patients with a CAC measurement. We also provide a score without inclusion of CAC for evaluation of the effect of including CAC in the novel risk score. We believe that the MESA risk score could be immediately used for communication of risk with patients after CAC scoring, to guide risk-based treatment decisions in clinical practice, as well as in designing future research studies that might use CAC to target high-risk subpopulations.

METHODS

STUDY PARTICIPANTS. MESA was designed to study the prevalence, risk factors, and progression of subclinical cardiovascular disease (CVD) in a multiethnic cohort. A detailed description of the study design and methods has been published previously (12). Briefly, 6,814 participants age 45 to 84 years who identified themselves as white, African-American, Hispanic, or Chinese were recruited from 6 U.S. communities from 2000 to 2002. All participants were free of clinically apparent CVD. The research was approved by the institutional review boards at all participating institutions, and all participants gave informed consent.

MEASUREMENT OF CAC. CAC was measured using electrocardiogram (ECG)-gated electron-beam computed tomography at 3 field centers and multidetector computed tomography at the other 3 field centers (12,13). Images were analyzed independently at a central reading center, and the amount of CAC was quantified using the Agatston scoring method (14). Rescan agreement was high using both electron-beam and multidetector computed tomography scanners (15). Interobserver and intraobserver agreement were also very high (Kappa = 0.93 and 0.90, respectively).

CORONARY HEART DISEASE ASCERTAINMENT. At intervals of 9 to 12 months, a telephone interviewer contacted each participant to inquire about interim hospitalizations, cardiovascular outpatient diagnoses and procedures, and deaths. Trained personnel abstracted medical records, and 2 physicians independently classified the events using pre-defined

Northwestern University Medical School, Chicago, Illinois; #Division of Cardiology, UCLA School of Medicine, Los Angeles, California; **Division of Epidemiology and Community Health, University of Minnesota School of Public Health, Minneapolis, Minnesota; ††Division of Cardiology, UT Southwestern Medical Center, Dallas, Texas; ‡‡Division of Clinical Sciences, UT Southwestern Medical Center, Dallas, Texas; §§University Clinic Essen, Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), Essen, Germany; ||||Institute of Medical Informatics, Biometry, and Epidemiology, University Clinic Essen, University of Duisburg, Essen, Germany; ¶¶Department of Radiology and Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee; and the ##Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina. Analysis and methods development for this project was supported by R01 HL 103729-01A1 from the National Heart, Lung, and Blood Institute. MESA was supported by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute, and by grants UL1-TR-000040, UL1 TR 001079, and UL1-RR-025005 from National Center for Research Resources. The HNR study is funded by a contract with the private Heinz Nixdorf Foundation and undergoes continuous monitoring by government agencies led by the Bundesministerium für Bildung und Forschung; and is also supported by the German Ministry of Education and Science. DHS was funded by the Donald W. Reynolds Foundation; and partially supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award Number UL1TR001105. Survey instrument development was funded in part by a Patient Care and Outcomes Research Grant from the American Heart Association. Dr. Budoff has received grant support from GE. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.
Listen to this manuscript's audio summary by *JACC* Editor-in-Chief Dr. Valentin Fuster.

Manuscript received April 28, 2015; revised manuscript received August 9, 2015, accepted August 10, 2015.

criteria. Hospital records were obtained for an estimated 98% of hospitalized cardiovascular events and some medical record-based information was available for 95% of outpatient encounters. All events through December 31, 2011, are included in this report.

Our endpoint consisted of incident hard CHD events: myocardial infarction (MI), resuscitated cardiac arrest, fatal CHD, and revascularization only if the participant also had prior or concurrent adjudicated angina. An MI required either abnormal cardiac biomarkers (2 times upper limits of normal) regardless of pain or ECG findings; evolving Q waves regardless of pain or biomarker findings; or a combination of chest pain, ST-T evolution or new left bundle branch block, and biomarker levels 1 to 2 times upper limits of normal. For suspected cardiovascular deaths based on International Classification of Diseases-10 underlying cause of death codes, a committee of MESA physicians classified CHD deaths using the death certificate, available medical records, and for out of hospital deaths, any next of kin interviews or physician questionnaires that could be obtained. A CHD death required a documented MI within the previous 28 days, chest pain within 72 h, or a history of CHD, and required the absence of a known noncardiac cause of death.

MEASUREMENT OF OTHER COVARIATES. Positive family history referred to a heart attack at any age in a parent, sibling, or child. The age at which the relative experienced the heart attack was not collected at baseline in MESA, precluding consideration of *premature* family history. Current smoking was defined as answering yes to the question "Have you smoked cigarettes during the last 30 days?" Resting blood pressure was measured 3 times in the seated position, and the average of the last 2 measurements was used in analysis. Medication use was determined by questionnaire. Additionally, participants were asked to bring containers for all medications used during the 2 weeks before the visit to the clinic.

STATISTICAL METHODS. Model development. Our list of candidate covariates included the traditional Framingham risk factors: age, sex, high-density lipoprotein cholesterol, total cholesterol, systolic blood pressure, antihypertensive medication use, current smoking, diabetes, and CAC. Additionally we considered family history of heart attack, lipid-lowering medication use, body mass index, and race/ethnicity. Interactions considered included: age, sex, race/ethnicity, and CAC with all other predictors; antihypertensive medications by systolic blood pressure; and lipid-lowering medications by total cholesterol. For continuous covariates, we explored nonlinearity by fitting generalized additive models with 4 degrees of freedom and adjusted for age, sex, and race/ethnicity (16). These models were fit using the "gam" package in Stata (17). There was evidence of potential nonlinearity for age and systolic blood pressure; hence, polynomial terms were considered. Although substantial nonlinearity was exhibited for untransformed CAC, the log(CAC + 1) transformation, which has been used previously (1), demonstrated no departures from linearity. It is critical to use CAC continuously to preserve all available information in the CAC variable, as has been preferred in prior CAC published data. The identical modeling strategy was used to develop the model that did not include CAC.

Shrinkage/penalization methods were used to avoid overfitting (18). We primarily used the method called Lasso ("least absolute shrinkage and selection operator"), which penalizes the sum of the absolute values of the regression coefficients (19,20). This leads to some coefficients shrinking all the way to 0, and hence, simultaneously performs variable selection. A penalized Cox proportional hazards model was used, and this was fit using the "penalized" package in R, version 3.1.1 (R Foundation for Statistical Computing, Vienna, Austria) (21). Shrinkage was done in 2 stages to allow inclusion of interaction terms only when the corresponding main effects were in the model. The first stage forced in the main effects of interest (unpenalized) and selected (via the Lasso) among the interaction and polynomial terms. The second stage started from the main effects plus selected interaction terms (if any) and then applied a ridge regression penalty. This penalty provides some shrinkage of the coefficients, but does not shrink any of the coefficients to 0. Selection of the tuning parameters at each stage was done via 10-fold cross-validated likelihood.

Proportional hazards for each variable were tested using Schoenfeld residuals in a standard Cox model that forced in all of the candidate main effects (22). There were no significant proportional hazards violations for any of the candidate main effects, and the global test was nonsignificant ($p = 0.33$).

External validation. The risk scores were validated in 2 independent longitudinal cohort studies, HNR and DHS. The risk score formula was sent to the coordinating centers of these studies, where the validation calculations were performed.

HNR is a single-center study that recruited a total of 4,814 Caucasians between 45 and 75 years of age from 3 neighboring cities in Germany between 2000 and 2003. Participants were a random sample derived from mandatory citizen registries. For this comparison of

the 2 study cohorts, we included 3,692 members of HNR who were free of clinical CVD at baseline and for whom complete covariate and follow-up data was available. Details about the design and recruitment strategy of the HNR have been previously published (23,24). Additionally, prior publications have described the comparability of the HNR and MESA cohorts with respect to CAC measurement, risk factors, and endpoints (25,26). Participants were followed for a median of 10.4 years (IQR: 9.2 to 11.3 years). For all primary study endpoints, hospital and nursing home records, including electrocardiograms, laboratory values, and pathology reports, were collected. For deceased subjects, death certificates were collected and interviews with general practitioners, relatives, and eyewitnesses were undertaken if possible. Medical records were obtained in 100% of all reported endpoints. An external endpoint committee blinded for risk factor status and CAC scores reviewed all documents and classified the endpoints.

The DHS is a multiethnic, population-based probability sample of Dallas County, Texas. The initial data collection was performed in 2000 to 2002 and included the collection of detailed socioeconomic, biomarker, and imaging data from each participant. Participants in DHS were age 30 to 65 years; however only 1,080 subjects age 45 to 65 years were included in this validation study. Details about the design of the DHS have been previously published (27). The DHS includes Caucasians, African Americans, and Hispanics. Participants were followed for a median of 9.3 years (IQR: 8.9 to 9.8 years). The DHS used a redundant strategy to ascertain clinical events. Death events were acquired for all DHS participants through the National Death Index, which was completed on December 31, 2010. Nonfatal events were captured through an annual detailed health survey administered by phone and by a unique data source, the Dallas-Fort Worth Hospital Council Data Initiative database, which captures hospital claims data for 77 hospitals in the metroplex area and represents >90% of the health care market volume in this region. More than 90% of participants from the CAC cohort were tracked using 1 or both of these mechanisms.

TABLE 1 Participant Characteristics—the MESA, HNR, and DHS Studies

	MESA (N = 6,726)	HNR (N = 3,692)	DHS (N = 1,080)
Age, yrs	62.1 ± 10.2	59.8 ± 7.7	52.7 ± 5.5
Male	3,176 (47.2)	1,714 (46.4)	466 (43.2)
Race/ethnicity			
Caucasian	2,622 (38.5)	3,692 (100)	409 (37.9)
Chinese American	803 (11.8)	–	–
African American	1,893 (27.8)	–	530 (49.1)
Hispanic American	1,496 (22.0)	–	122 (11.3)
Other	–	–	19 (1.8)
Diabetes	859 (12.7)	470 (12.7)	148 (13.7)
Current smoker	887 (13.0)	831 (22.5)	272 (25.2)
Total cholesterol, mg/dl	194.2 ± 35.7	231.0 ± 38.9	189.3 ± 41.5
HDL cholesterol, mg/dl	51.0 ± 14.8	58.8 ± 17.0	51.4 ± 14.9
Lipid-lowering medications	1,100 (16.2)	368 (10.0)	104 (9.6)
Systolic blood pressure, mm Hg	126.6 ± 21.5	132.9 ± 20.6	128.5 ± 19.5
Antihypertensive medications	2,536 (37.2)	1,241 (33.6)	330 (30.6)
Family history of heart attack			
No	3,611 (53.7)	2,671 (72.3)	436 (40.4)
Yes	2,699 (40.1)	1,021 (27.7)	501 (46.4)
unknown	416 (6.2)	–	143 (13.2)
Coronary artery calcium (Agatston)			
0	3,416 (50.1)	1,138 (30.8)	361 (33.4)
1–99	1,794 (26.3)	1,512 (40.9)	565 (52.3)
100–399	927 (13.6)	659 (17.9)	113 (10.5)
400+	677 (9.9)	383 (10.4)	41 (3.8)
10-yr (Kaplan-Meier) CHD rate	6.5%	7.5%	5.5%

Values are mean ± SD or n (%). For DHS, the "Other" race/ethnic category includes the following: American Indian/Pacific Islander/Alaskan Native/Asian/East Indian.

CHD = coronary heart disease; DHS = Dallas Heart Study; HDL = high-density lipoprotein; HNR = Heinz Nixdorf Recall Study; MESA = Multi-Ethnic Study of Atherosclerosis.

Performance metrics. We assessed 2 aspects of model performance: discrimination and calibration. Discrimination was assessed using Harrell's C-statistic, the discrimination slope (the absolute difference between the average predicted 10-year probability of an event for those who experienced an event minus the average for those who did not), and the area under the survival receiver-operator characteristic (ROC) curve (18,28,29). Calibration was assessed via the calibration slope (calculated as the slope of the linear regression of the event indicator on the predicted probability) and calibration-in-the-large (the difference between the observed event rate and the average predicted 10-year event rate) (18).

Presentation. A computerized version of the risk score will be posted on the MESA Website for ease of use by clinicians, patients, and researchers. This MESA risk score application requires only the input of the traditional risk factors and the CAC score and is the preferred way for interested parties to use the risk score.

RESULTS

Participants in MESA were followed for a median of 10.2 years (IQR: 9.7 to 10.7 years), and 422 CHD events were observed, with first events including 68 CHD deaths, 190 nonfatal MIs, 149 angina-driven revascularizations, and 15 resuscitated cardiac arrests. A total of 88 participants were excluded from this study, (5 participants were found to have a

TABLE 2 MESA 10-Year CHD Risk Prediction Models

	Risk Factors Only			Risk Factors and CAC		
	Hazards Ratio	Beta Coefficient	p Value	Hazards Ratio	Beta Coefficient	p Value
Age, yrs	1.05	0.0455	<0.0001	1.02	0.0172	.007
Male	2.12	0.7496	<0.0001	1.50	0.4079	<0.001
Race/ethnicity						
Non-Hispanic white	Ref	0	–	Ref	0	–
Chinese American	0.60	−0.5055	<0.01	0.71	−0.3475	0.07
African American	0.81	−0.2111	0.066	1.04	0.0353	0.70
Hispanic	0.83	−0.1900	0.11	0.98	−0.0222	0.88
Diabetes	1.68	0.5168	<0.0001	1.48	0.3892	0.002
Current smoker	1.61	0.4732	<0.001	1.45	0.3717	0.005
Total cholesterol, mg/dl	1.01	0.0053	<0.0001	1.00	0.0043	<0.001
HDL cholesterol, mg/dl	0.99	−0.0140	<0.001	0.99	−0.0114	0.003
Lipid-lowering meds	1.28	0.2473	0.003	1.13	0.1206	0.32
Systolic blood pressure, mm Hg	1.01	0.0085	0.0002	1.01	0.0066	0.004
Antihypertensive meds	1.40	0.3381	0.0013	1.26	0.2278	0.033
Family history of heart attack	1.57	0.4522	<0.0001	1.38	0.3239	<0.001
ln (CAC + 1)	NA	NA	NA	1.32	0.2743	<0.0001
Baseline survival at 10 yrs, S(10)	0.99963			0.99833		

The p values are based on a standard Cox proportional hazards model. To estimate the 10-year risk of a CHD event for a particular person, multiply the values of the risk factors by the corresponding beta coefficient and sum these quantities up to yield a value (call it A following the notation in Wilson et al. [38]). Then calculate: B = exp(A). Finally, the 10-year risk is given by $1 - S(10)^B$. Alternatively, see the online calculator on the MESA Website.

CAC = coronary artery calcium; other abbreviations as in Table 1.

pre-baseline event, 28 were without follow-up, and 55 were missing covariate data). The remaining sample size was 6,726.

Table 1 provides the baseline characteristics of the MESA cohort and the 2 validation cohorts. All 3 cohorts are sex balanced. DHS participants are 10 years younger on average than participants in MESA and HNR. HNR is exclusively Caucasian, whereas DHS includes African Americans and Hispanic Americans, similar to MESA. Rates of diabetes are similar across cohorts. MESA includes fewer current smokers than either HNR or DHS. MESA and DHS are similar in terms of lipid levels, blood pressure, and family history of heart attack. HNR participants have worse lipid profiles and blood pressure levels, but less positive family history of heart attack. Kaplan-Meier 10-year rates of CHD were highest for HNR (7.5%), followed by MESA (6.5%), and finally DHS (5.5%).

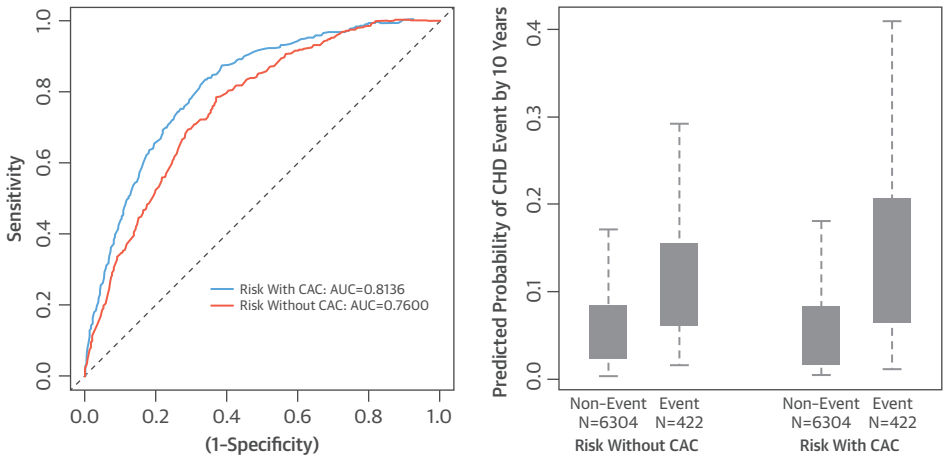
Table 2 provides the estimated hazards ratios, coefficients, and baseline hazards for a 10-year CHD risk prediction model with and without CAC. The table provides the baseline survival function and risk factor coefficients that have been incorporated into the online MESA risk score application. No interaction or polynomial terms were retained in the model, nor were body mass index or diastolic blood pressure included.

Figure 1 illustrates the internal discrimination properties of our 2 models. The area under the survival ROC curve within MESA was 0.81 for the model with CAC, indicating excellent discrimination between events and nonevents. Comparison with our internally developed risk score without CAC, which has an area under the survival ROC of 0.76, provides evidence of the significant improvement due to inclusion of CAC ($p < 0.0001$). The boxplot in the second panel of Figure 1 shows separation in predicted 10-year risk between events and nonevents, and this is also improved by the addition of CAC. The difference between events and nonevents was significant for each version of the score ($p < 0.001$ for both). The discrimination slope was 0.086 for the score with CAC. That is, those who experienced events had an average predicted 10-year risk that was 8.6% higher than those who did not. In contrast, the discrimination slope was 0.052 for the score without CAC, or an average separation of 5.2% higher predicted risk for those with events.

We also evaluated our internal discrimination performance in various subsets of the MESA cohort. The area under the survival ROC was better for non-Caucasians relative to Caucasians (Chinese: 0.85, Black: 0.80, Hispanic: 0.86, Caucasian: 0.79), better for young (under 65 years: 0.82) than old (≥65 years: 0.76), and women (0.81) compared with men (0.79).

External validation in both the HNR and DHS studies is shown in Table 3 and the Central Illustration, and provides evidence for very good to

FIGURE 1 Discrimination of the MESA CHD Risk Score Within the Development Cohort



(Left) Receiver-operator characteristic (ROC) curves for the risk scores with and without coronary artery calcium (CAC) applied within the MESA (Multi-Ethnic Study of Atherosclerosis) cohort. **(Right)** A boxplot of the predicted 10-year coronary heart disease (CHD) probabilities for each score stratified by event status. The **shaded box** covers the middle 50% of the data. AUC = area under the curve.

excellent discrimination and calibration for the model including CAC. Harrell's C-statistic was 0.779 in the HNR study and 0.816 in the DHS. There was a 7.8% to 9.5% difference in predicted probability between events and nonevents. Observed and predicted risks were close across the range of the score, as seen in the **Central Illustration**, with the exception of some underestimation seen in the highest risk group of DHS. There was no evidence of poor calibration based on the Hosmer-Lemeshow goodness-of-fit statistics ($p > 0.22$ for each cohort). Externally validated calibration slopes were 0.90 for HNR and 1.19 for DHS (perfect calibration yields a slope of 1.0). Mean calibration, or calibration-in-the-large, was excellent for both studies (−0.50% for HNR and −0.46% for DHS), indicating that on average, predicted risk was within one-half of a percent of the observed event rate. In comparison, the model without CAC also was not as well calibrated in HNR (calibration slope 0.74), and discrimination was substantially worse (C-statistic = 0.720, discrimination slope 0.053). In DHS, the C-statistic for the model without CAC was very good (C-statistic = 0.782) but the discrimination slope was only 0.046, and the calibration slope was large at 1.55.

Figure 2 displays a sample patient case. A 70-year-old Hispanic man with mild treated hypertension and no other traditional risk factors would have a 10-year CHD risk of 9.3% before considering CAC data and 3.1% after a CAC score result of 0.

TABLE 3 Evaluation of Model Performance

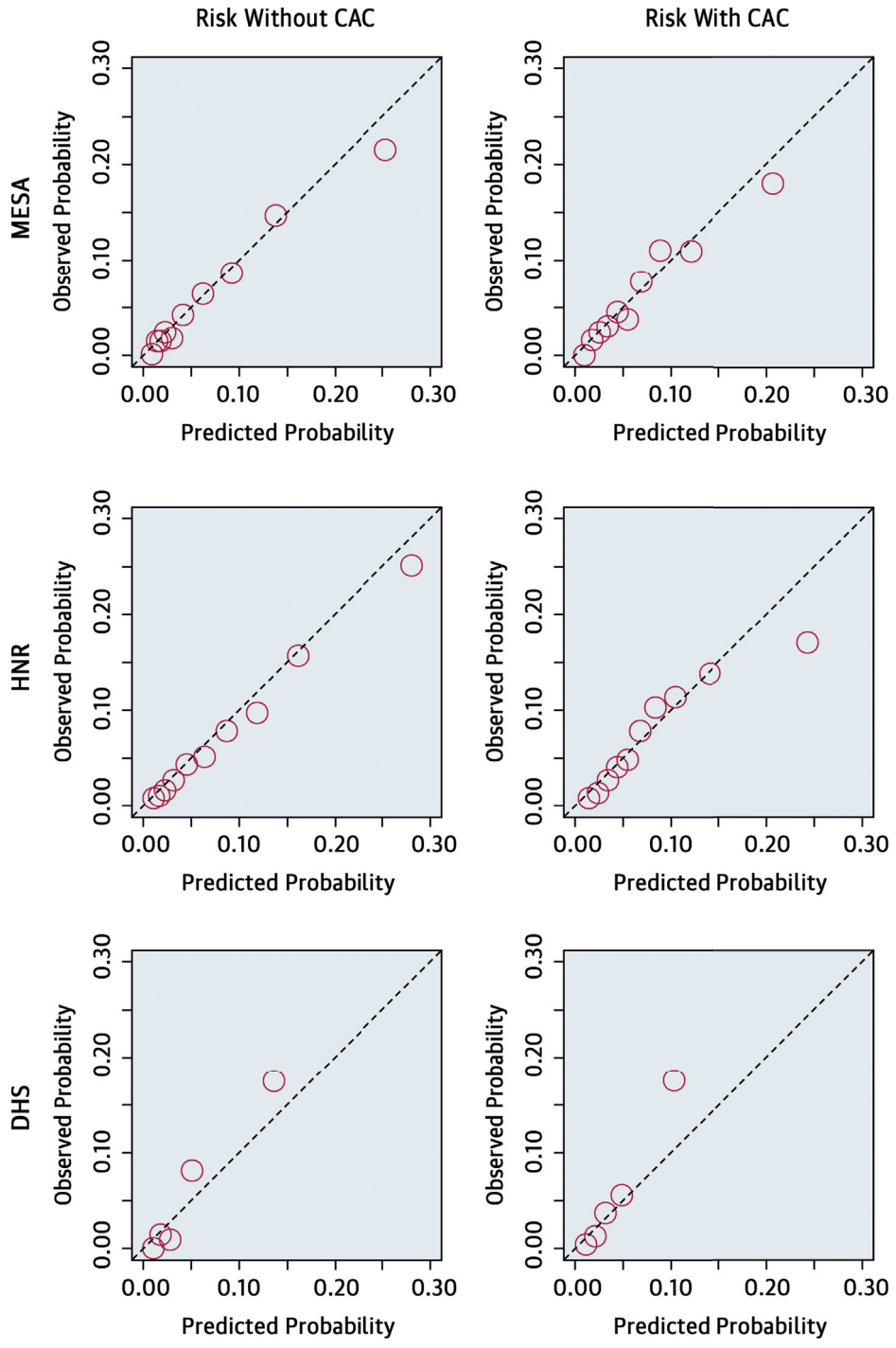
	MESA	HNR	DHS
Sample size	6,726	3,692	1,080
CHD events, n	422	274	58
Model with risk factors only			
Harrell's C-statistic	0.750	0.720	0.782
Discrimination slope	0.052	0.053	0.046
Calibration slope	0.834	0.740	1.55
Model with risk factors and CAC			
Harrell's C-statistic	0.800	0.779	0.816
Discrimination slope	0.086	0.095	0.078
Calibration slope	0.857	0.899	1.19

Abbreviations as in **Tables 1 and 2**.

DISCUSSION

CAC is a direct measurement of 1 component of atherosclerotic plaque in the coronary arteries, and a potent predictor of future CHD events (1–11). Risk prediction equations are recommended for clinical use to select the best candidates for preventive therapies such as cholesterol-lowering medications (30). One commonly stated limitation for clinical CAC scoring is the absence of a risk calculator for integrating this information into global cardiovascular risk assessment (31). Here, we present a predictive algorithm to integrate CAC measurement with traditional risk factors and demonstrate that a risk score that includes CAC improves CHD risk prediction compared with single measurements of traditional

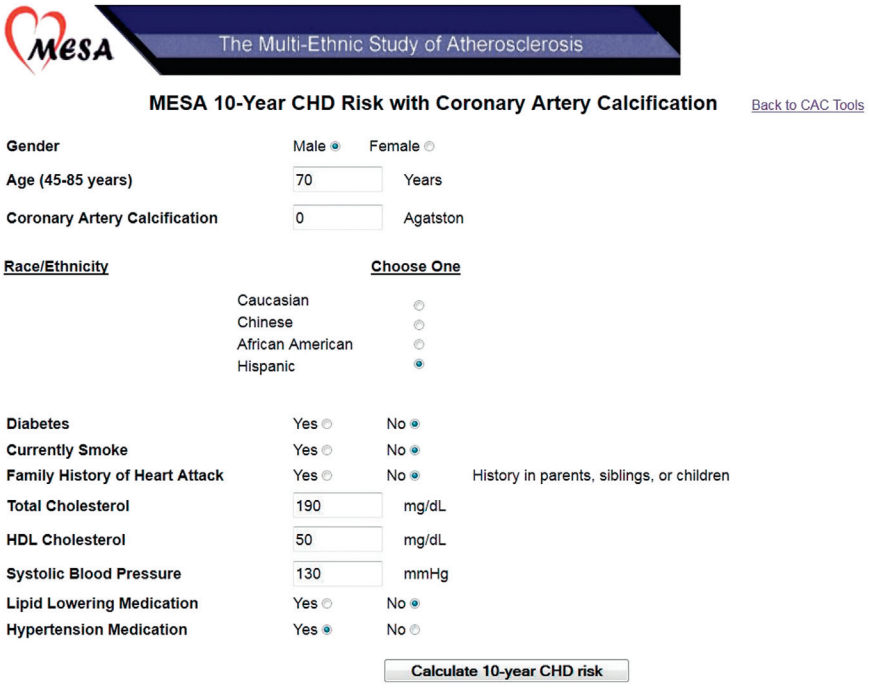
CENTRAL ILLUSTRATION **MESA CHD Risk Score Using CAC: Calibration of the MESA CHD Risk Score**



McClelland, R.L. et al. J Am Coll Cardiol. 2015; 66(15):1643–53.

The observed versus the predicted event rates are presented. Predicted risks were divided into 10 equal sized bins for the MESA (Multi-Ethnic Study of Atherosclerosis) and HNR (Heinz Nixdorf Recall Study). For the DHS (Dallas Heart Study), only 5 bins were used due to the smaller sample size. CAC = coronary artery calcium; CHD = coronary heart disease.

FIGURE 2 Online Risk Score Calculator



A screenshot of the online risk score calculator on the MESA Website. Here we see the calculator being used for an older man with fairly favorable risk factor profile but 0 CAC. Under a risk score without CAC, 10-year estimated CHD risk is 9.3%, due in large part to his age. Once we factor in that he has no detectable CAC, estimated risk is only 3.1%. HDL = high-density lipoprotein cholesterol; other abbreviations as in Figure 1.

risk factors alone (32). The use of this equation can be considered as part of the "risk discussion" between a clinician and patient when CAC imaging has been performed (33).

We also show that the algorithm generalizes well to 2 external populations. These validation cohorts included 1 large study in Germany (HNR) that follows very similar protocols to MESA and has similar age and risk factor distribution. The second validation cohort, the DHS, is U.S.-derived and multiethnic, similar to MESA and the U.S. population. Our algorithm includes a term for race/ethnicity; however, consistent with an earlier report from MESA by Detrano et al. (1), we did not find that the associations of CAC or other risk factors with CHD events differed by race/ethnicity. Our results are consistent with prior studies demonstrating that CAC improves discrimination of CHD events in the MESA population and the HNR study (2,4–11).

We evaluated 2 important aspects of a risk prediction algorithm: discrimination and calibration. Discrimination refers to the ability of the risk score to separate those who ultimately have events from those who do not. The statistics that reflect this are the C-statistic and the discrimination slope. C-statistics for existing risk scores, including the Framingham CHD risk score and the 2013 American College of Cardiology (ACC)/American Heart Association (AHA) risk estimator, evaluated in the MESA data fall in the range of 0.65 to 0.75 depending on the subset being investigated (34). Improvement to the 0.78 to 0.81 range, as demonstrated in our 2 external validations, represents a substantial gain in terms of this metric. We also note that comparing discrimination with a recalibrated AHA/ACC risk score (30) (C-index 0.71) to a strategy of adding CAC to the AHA/ACC score (C-index 0.78) to our risk score (C-index 0.80) supports the strategy of deriving a new score to incorporate CAC, rather than attempting to simply add CAC onto the existing risk score.

The discrimination slope represents the separation between events and nonevents in terms of

average predicted risk. Results indicate that those who experience events will have predicted 10-year risk estimates that are on average 8% to 9% higher than the nonevent group. Importantly, the MESA CHD risk score with CAC offers a substantial improvement in separation. Finally, calibration refers to the agreement between observed and predicted event rates in the population. The HNR validation results indicate excellent calibration, and the DHS results indicate very good calibration. This is important, as the AHA/ACC risk score is known to overestimate risk and exhibit poor calibration in MESA (34).

Traditional CHD risk scores are strongly influenced by age, sex, and race/ethnicity. Importantly, including CAC in the risk score markedly decreases the effect of these demographic risk factors (**Table 2**). This is likely explained by the fact that CAC serves to integrate the effect of all measured (and unmeasured) risk factors over the course of an individual's lifetime up until the point of CAC measurement. This is important in terms of individualization of risk, and avoids scenarios where, for instance, all men over a particular age are deemed high risk based on their chronological age alone.

We also considered a direct comparison with existing risk scores; however, several factors limit our enthusiasm for this comparison. Existing risk scores have been shown to be poorly calibrated and/or have low discrimination in MESA (34), and thus, comparison of any new score developed in MESA (even in the absence of CAC) is guaranteed to show improvement. It would be unclear if this improvement is the result of an improved score, the addition of CAC, a problem with the original score, or some combination of these and other factors. In other words, comparing to an existing score would likely overstate the added value due to inclusion of CAC. For this reason, we opted to make our comparison with a MESA version of a risk score, developed using the same modeling strategy as the CAC enhanced score—improvements in prediction of CHD events over this score are clearly due to the addition of CAC to the risk scoring paradigm. Our goal here was not to develop a new score with traditional risk factors only, but to provide an appropriate baseline for evaluation of our new score that incorporates CAC.

We included family history because it is a strong independent predictor of risk and is easily obtainable with no additional testing. We opted to include CAC as the sole marker of subclinical disease as it has been shown to offer the largest incremental prediction improvement over traditional risk factors. There may be potential utility in exploring whether other subclinical measures (i.e., ankle-brachial index, carotid plaque, ECG abnormalities) offer further incremental improvement in future MESA risk scores modeling stroke or a more inclusive CVD outcome.

Existing risk scores differ in their choice of endpoint. We modeled a CHD endpoint, similar to the Framingham risk score as described in the Adult Treatment Panel III (ATP 3) report (35), but dissimilar to other risk scores that also include stroke (30,36,37), angina (38,39), and other events such as peripheral vascular disease, transient ischemic attack, and heart failure (39). We opted against using a composite CVD endpoint, because each CVD component has different associations with risk factors. In addition, CAC has been shown to be much more strongly associated with CHD events than with cerebrovascular disease (40) and is more likely to be used clinically for CHD risk prediction.

The algorithm is a prediction tool, and the terms in the model should not be interpreted causally. For example, the term for antihypertensive medications denotes an increased risk for those on medications. This reflects the fact that those with treated hypertension tend to be a higher-risk population, not that the medications themselves increase their risk. The term captures a *combination* of the increased risk of this population and the beneficial effect the medication has for those taking it. Similar reasoning holds for the effect of lipid-lowering therapy. An additional cautionary note is that the risk estimates will have considerable variability in subsets of participants that are rare in the development data. For instance, participants with very high CAC (>400 or >1,000) despite having a "normal" risk factor profile are rare in MESA. For participants with more "typical" risk factor profiles, the accuracy of the risk estimate will be optimized.

The MESA study includes many participants that were taking blood pressure or lipid-lowering medication at the baseline examination. A strength of the MESA risk score is that it remains relevant for these commonly encountered patients, many of whom are treated for hypertension and/or hypercholesterolemia at the time of the initial encounter. For example, the risk scores can be used by the physician to motivate patient life-style change, encourage adherence to existing therapy, or to guide decisions about treatment intensity. The MESA risk score may be valuable for guiding decisions about the net benefit of daily aspirin therapy in primary prevention (41). When making treatment initiation decisions, the value of lipid-lowering or antihypertensive medications can simply be entered as "zero." This allows the risk score to be useful in more situations.

STUDY STRENGTHS AND LIMITATIONS. Strengths of this study include the large, modern, community-based multiethnic cohort and the use of statistical techniques to provide a model that performs well when applied outside of the development cohort. Independent validation of the model in 2 contemporary cohorts—1 international from Germany and 1 U.S.-based multiethnic study—provides evidence of external validity. The age range of MESA was 45 to 84 years at study baseline, and hence, the algorithm is limited to this range. Although MESA is multiethnic, there are many race/ethnicities not represented in the study, and the utility of the algorithm in these groups is unknown. Limitations of this study exist. The development cohort consisted of individuals aged 45 to 85 and in 1 of 4 race/ethnic groups. The generalizability of the score outside of these demographics is unknown. Additionally, our external validation cohorts did not include Chinese participants.

CONCLUSIONS

The MESA CHD risk score is the first available algorithm incorporating CAC with traditional risk factors for 10-year risk prediction. In addition to its use in direct clinical encounters, the MESA risk score can be used by radiologists and cardiologists when interpreting and reporting CAC scores. Similar to the current practice of reporting CAC percentiles or "arterial age" to place CAC results into clinical context, scan readers can now calculate and provide a "post-test" 10-year CHD risk after CAC scanning based on the MESA risk score. This updated 10-year risk could be used to help make therapeutic decisions, such as the decision to start statin or aspirin therapy in primary prevention. Future guidelines from the Society of Cardiovascular Computed Tomography might consider recommending this practice in routine CAC score reporting. Additionally, future iterations of U.S. and international prevention guidelines may consider use of the MESA risk score as an alternative risk score to existing algorithms when CAC score results are available.

ACKNOWLEDGMENTS The authors thank the other investigators, the staff, and the participants of the MESA, HNR, and DHS studies for their valuable contributions.

REPRINT REQUESTS AND CORRESPONDENCE: Dr. Robyn L. McClelland, Department of Biostatistics, University of Washington, Building 29, Suite 310, 6200 Northeast 74th Street, Seattle, Washington 98115. E-mail: rmcclell@u.washington.edu.

PERSPECTIVES

COMPETENCY IN PATIENT CARE AND PROCEDURAL SKILLS: Detection of CAC by computed tomography imaging can add to traditional risk factors for prediction of ischemic risk, but no validated method other than the MESA risk score has been available to incorporate the CAC score in estimating an individual's 10-year risk of coronary events.

TRANSLATIONAL OUTLOOK: Prospective studies are needed to validate the MESA risk score in independent cohorts, enhance the clinical application of CAC scoring in cardiovascular risk assessment, and inform the integration of this approach in future prevention guidelines.

REFERENCES

1. Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N Engl J Med 2008;358:1336–45.

2. Erbel R, Möhlenkamp S, Moebus S, et al. Coronary risk stratification, discrimination, and reclassification improvement based on quantification of subclinical coronary atherosclerosis: The Heinz Nixdorf Recall Study. J Am Coll Cardiol 2010;56:1397–406.

3. Paixao ARM, Berry JD, Neeland IJ, et al. Coronary artery calcification and family history of myocardial infarction in the Dallas Heart Study. J Am Coll Cardiol Img 2014;7:679–86.

4. Polonsky TS, McClelland RL, Jorgensen NW, et al. Coronary artery calcium score and risk classification for coronary heart disease prediction. JAMA 2010;303:1610–6.

5. Yeboah J, McClelland RL, Polonsky TS, et al. Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate risk individuals: the Multi-Ethnic Study of Atherosclerosis. JAMA 2012;308:788–95.

6. Jain A, McClelland RL, Polak JF, et al. Cardiovascular imaging for assessing cardiovascular risk in asymptomatic men versus women: the Multi-Ethnic Study of Atherosclerosis (MESA). Circ Cardiovasc Imaging 2010;4:8–15.

7. Simon A, Megnien JL, Chironi G. The value of carotid intima-media thickness for predicting cardiovascular risk. Arterioscler Thromb Vasc Biol 2010;30:182–5.

8. Peters SA, Bakker M, den Ruijter HM, Bots ML. Added value of CAC in risk stratification for cardiovascular events: a systematic review. Eur J Clin Invest 2012;42:110–6.

9. Peters SA, den Ruijter HM, Bots ML, Moons KG. Improvements in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. Heart 2012;98:177–84.

10. Möhlenkamp S, Lehmann N, Moebus S, et al., for the Heinz Nixdorf Recall Study Investigators. Quantification of coronary atherosclerosis and inflammation to predict coronary events and all-cause mortality. J Am Coll Cardiol 2011;57:1455–64.

11. Erbel R, Lehmann N, Möhlenkamp S, et al., for the Heinz Nixdorf Recall Study Investigators. Subclinical coronary atherosclerosis predicts cardiovascular risk in different stages of hypertension: result of the Heinz Nixdorf Recall Study. Hypertension 2012;59:44–53.

12. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. Am J Epidemiol 2002;156:871–81.

13. Carr JJ, Nelson JC, Wong ND, et al. Calcified coronary artery plaque measurement with cardiac CT in population-based studies: standardized

protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) Study. Radiology 2005; 234:35–43.

14. Agatston AS, Janowitz WR, Hildner FJ, et al. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol 1990;15:827–32.

15. Detrano RC, Anderson M, Nelson J, et al. Coronary calcium measurements: effect of CT scanner type and calcium measure on rescan reproducibility—MESA Study. Radiology 2005;236:477–84.

16. Hastie TJ, Tibshirani R. Generalized Additive Models. London: Chapman and Hall, 1990.

17. Royston P, Ambler G. Generalized additive models. Stata Tech Bull 1998;42:38–43.

18. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer, 2010.

19. Tibshirani R. Regression shrinkage and selection via the lasso. J Royal Statist Soc B 1996;58: 267–88.

20. Tibshirani R. The LASSO method for variable selection in the Cox model. Stat Med 1997;16: 385–95.

21. Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. Biom J 2010;52:70–84.

22. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239–41.

23. Schmermund A, Möhlenkamp S, Stang A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Am Heart J 2002;144:212–8.

24. Stang A, Moebus S, Dragano N, et al., for the Heinz Nixdorf Recall Study Investigation Group. Baseline recruitment and analyses of nonresponse of the Heinz Nixdorf Recall Study: identifiability of phone numbers as the major determinant of response. Eur J Epidemiol 2005;20:489–96.

25. Erbel R, Delaney JA, Lehmann N, et al. Signs of subclinical coronary atherosclerosis in relation to risk factor distribution in the Multi-Ethnic Study of Atherosclerosis (MESA) and the Heinz Nixdorf Recall Study (HNR). Eur Heart J 2008;29:2782–91.

26. Budoff MJ, Möhlenkamp S, McClelland R, et al. A comparison of outcomes with coronary artery calcium scanning in unselected populations: the Multi-Ethnic Study of Atherosclerosis (MESA) and Heinz Nixdorf RECALL study (HNR). J Cardiovasc Comput Tomogr 2013;7:182–91.

27. Victor RG, Haley RW, Willett DL, et al. The Dallas Heart Study: a population-based probability sample for the multidisciplinary study of ethnic differences in cardiovascular health. Am J Cardiol 2004;93:1473–80.

28. Harrell FE Jr., Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. JAMA 1982; 247:2543–6.

29. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 2000;56: 337–44.

30. Goff DC Jr., Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63:2935–59.

31. Anderson C, Vasan RS. Is there a role for coronary artery calcium scoring for management of asymptomatic patients at risk for coronary artery disease? Clinical risk scores are sufficient to define primary prevention treatment strategies among asymptomatic patients. Circ Cardiovasc Imaging 2014;7:390–7.

32. Blaha MJ, Silverman MG, Budoff MJ. Is there a role for coronary artery calcium scoring for management of asymptomatic patients at risk for coronary artery disease? Clinical risk scores are not sufficient to define primary prevention treatment strategies among asymptomatic patients. Circ Cardiovasc Imaging 2014;7:398–408.

33. Martin SS, Sperling LS, Blaha MJ, et al. Clinician-patient risk discussion for atherosclerotic cardiovascular disease prevention: importance to implementation of the 2013 ACC/AHA guidelines. J Am Coll Cardiol 2015;65:1361–8.

34. DeFilippis AP, Young R, Carrubba CJ, et al. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. Ann Intern Med 2015;162: 266–75.

35. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). JAMA 2001;285: 2486–97.

36. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA 2007;297:611–9.

37. Ridker PM, Paynter NP, Rifai N, et al. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. Circulation 2008;118: 2243–51.

38. Wilson PF, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837–47.

39. D'Agostino RB Sr., Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008;117:743–53.

40. Folsom AR, Kronmal RA, Detrano RC, et al. Coronary artery calcification compared with carotid intima-media thickness in prediction of cardiovascular disease incidence: The Multi-Ethnic Study of Atherosclerosis (MESA). Arch Intern Med 2008;168:1333–9.

41. Miedema MD, Duprez DA, Misialek JR, et al. Use of coronary artery calcium testing to guide aspirin utilization for primary prevention: estimates from the Multi-Ethnic Study of Atherosclerosis. Circ Cardiovasc Qual Outcomes 2014;7: 453–60.

KEY WORDS atherosclerosis, coronary disease, epidemiology, risk prediction